

What is claimed is:

## CLAIMS

- 1     1.     A method for a storage operating system implemented in a storage system to con-  
2     currently perform readahead operations for a plurality of different read streams estab-  
3     lished in one or more files, directories, vdisks or luns stored in the storage system, the  
4     method comprising:  
5         receiving a client read request at the storage system, the client read request indi-  
6     cating client-requested data for the storage operating system to retrieve from a file, di-  
7     rectory, vdisk or lun stored in the storage system;  
8         determining whether the received client read request matches any of a plurality of  
9     readset data structures ("readsets") allocated for the file, directory, vdisk or lun contain-  
10    ing the client-requested data; and  
11         performing readahead operations in accordance with a set of readahead metadata  
12    stored in a readset that is determined to match the received client read request.
- 1     2.     The method of claim 1, further comprising:  
2         allocating at least one readset for each of the one or more files, directories, vdisks  
3     or luns in which the plurality of different read streams is established;  
4         generating a separate set of readahead metadata for each of the plurality of differ-  
5     ent read streams; and  
6         storing each generated set of readahead metadata in a different readset allocated  
7     for the file, directory, vdisk or lun in which the read stream associated with the generated  
8     set of readahead metadata is established.
- 1     3.     The method of claim 1, further comprising:  
2         initializing each allocated readset to store a predetermined set of values.
- 1     4.     The method of claim 2, wherein the number of readsets allocated for a file, direc-  
2     tory, vdisk or lun depends on the size of that file, directory, vdisk or lun.

- 1     5.     The method of claim 4, wherein the number of readsets allocated for a file, direc-  
2     tory, vdisk or lun is dynamically increased as the size of that file, directory, vdisk or lun  
3     is increased.
  
- 1     6.     The method of claim 1, wherein a first readset is determined to match the received  
2     client read request if the first readset stores a set of readahead metadata associated with a  
3     read stream that is extended by the client-requested data.
  
- 1     7.     The method of claim 1, wherein a second readset is determined to match the re-  
2     ceived client read request when the client-requested data is located within a predeter-  
3     mined fuzzy range associated with the second readset.
  
- 1     8.     The method of claim 7, wherein the fuzzy range is derived based on a multiple of  
2     a number of client-requested data blocks specified in the received client read request.
  
- 1     9.     The method of claim 7, wherein the fuzzy range extends in both a forward direc-  
2     tion and a backward direction in relation to a last data block retrieved in a read stream  
3     associated with the second readset.
  
- 1     10.    The method of claim 1, wherein a third readset is determined to match the re-  
2     ceived client read request if the third readset is determined to be unused.
  
- 1     11.    The method of claim 10, wherein the third readset is determined to be unused  
2     when a level value stored in the third readset equals a special indicator value.
  
- 1     12.    The method of claim 1, wherein readahead operations are not performed if the  
2     storage operating system determines that the file, directory, vdisk or lun containing the  
3     client-requested data is accessed using a random access style.

- 1 13. The method of claim 12, wherein a DAFS cache hint included in the received cli-  
2 ent read request indicates that the file, directory, vdisk or lun containing the client-  
3 requested data is accessed using a random access style.
- 1 14. The method of claim 1, wherein readahead operations are not performed unless:  
2 (i) a readset is determined to match the received client read request; and  
3 (ii) the matching readset stores a set of readahead metadata associated  
4 with a read stream that is extended by the client-requested data past a predeter-  
5 mined data block or memory address.
- 1 15. The method of claim 1, further comprising:  
2 if the received client read request does not match any of the readsets allocated for  
3 the file, directory, vdisk or lun containing the client-requested data, then performing the  
4 steps:  
5 identifying the received client read request as being the first read  
6 request in a new read stream;  
7 generating a set of readahead metadata associated with the new  
8 read stream;  
9 selecting for reuse one of the readsets allocated for the file, direc-  
10 tory, vdisk or lun containing the client-requested data; and  
11 storing the generated set of readahead metadata associated with the  
12 new read stream in the readset selected for reuse.
- 1 16. The method of claim 15, wherein the readset selected for reuse stores a level  
2 value that is less than or equal to level values stored in each of the other readsets associ-  
3 ated with the file, directory, vdisk or lun containing the client-requested data.
- 1 17. The method of claim 1, wherein the client read request received at the storage  
2 system is a file-based client read request.

1 18. The method of claim 1, wherein the client read request received at the storage  
2 system is a block-based client read request.

1 19. A multiprotocol storage appliance, comprising:  
2 a memory configured to store instructions for implementing a storage operating  
3 system and further configured to store at least one readset data structure ("readset") asso-  
4 ciated with a file, directory, vdisk or lun, the readset including:  
5 a level value indicating the readset's relative age with respect to other  
6 readsets implemented in the memory, the level value being equal to a value that is  
7 between a predetermined upper-bound level value and a predetermined lower-  
8 bound level value, inclusive;  
9 a count value indicating the number of client read requests processed in a  
10 read stream associated with the readset;  
11 a next readahead value indicating a predetermined file offset or memory  
12 address, wherein readahead operations are performed by the storage operating  
13 system when the read stream associated with the readset is extended past the pre-  
14 determined file offset or memory address; and  
15 a readahead size value indicating the amount of data that is prefetched by  
16 the storage operating system when readahead operations are performed for the  
17 read stream associated with the readset.

1 20. The multiprotocol storage appliance of claim 19, further comprising:  
2 a processor;  
3 at least one storage device configured to store one or more files, directories,  
4 vdisks or luns; and  
5 a network adapter that is adapted to receive at least one client read request, the  
6 client read request indicating client-requested data contained in a file, directory, vdisk or  
7 lun stored in the at least one storage device.

- 1 21. The multiprotocol storage appliance of claim 19, wherein the count value is ini-  
2 tially set equal to zero.
- 1 22. The multiprotocol storage appliance of claim 19, wherein the level value is set  
2 equal to a special indicator value when the readset is first allocated, wherein the special  
3 indicator value may be outside the range of values defined by the predetermined upper-  
4 bound level value and predetermined lower-bound level value.
- 1 23. The multiprotocol storage appliance of claim 19, wherein the level value is in-  
2 cremented by a first step size when the received client read request is determined to  
3 match the readset, and the level value is decremented by a second step size when the re-  
4 ceived client read request causes the storage operating system to reuse a different readset  
5 data structure in the memory.
- 1 24. The multiprotocol storage appliance of claim 23, wherein the first and second step  
2 sizes equal one.
- 1 25. The multiprotocol storage appliance of claim 23, wherein if the level value is less  
2 than a predetermined initial level value and the level value is to be incremented, the level  
3 value is set equal to the predetermined initial level value.
- 1 26. The multiprotocol storage appliance of claim 19, wherein the count value is in-  
2 cremented by one if the received client read request is determined to match the readset.
- 1 27. The multiprotocol storage appliance of claim 20, wherein the readset is deter-  
2 mined to be eligible for reuse if (i) the level value stored in the readset is the lowest level  
3 value among all the readsets associated with the file, directory, vdisk or lun containing  
4 the client-requested data or (ii) the readset is the most-recently accessed readset associ-  
5 ated with the file, directory, vdisk or lun containing the client-requested data and the  
6 readset stores a count value equal to one.

1 28. The multiprotocol storage appliance of claim 27, wherein the most-recently ac-  
2 cessed readset associated with the file, directory, vdisk or lun containing the client-  
3 requested data is located at the head of a linked list of readsets associated with the file,  
4 directory, vdisk or lun containing the client-requested data.

1 29. A storage system that employs a storage operating system to concurrently perform  
2 readahead operations for a plurality of different read streams established in one or more  
3 files, directories, vdisks or luns stored in the storage system, the method comprising:  
4 means for receiving a client read request at the storage system, the client read re-  
5 quest indicating client-requested data for the storage operating system to retrieve from a  
6 file, directory, vdisk or lun stored in the storage system;  
7 means for determining whether the received client read request matches any of a  
8 plurality of readset data structures ("readsets") allocated for the file, directory, vdisk or  
9 lun containing the client-requested data; and  
10 means for performing readahead operations in accordance with a set of readahead  
11 metadata stored in a readset that is determined to match the received client read request.

1 30. A computer-readable media comprising instructions for execution in a processor  
2 for the practice of a method for a storage operating system implemented in a storage  
3 system to concurrently perform readahead operations for a plurality of different read  
4 streams established in one or more files, directories, vdisks or luns stored in the storage  
5 system, the method comprising:  
6 receiving a client read request at the storage system, the client read request indi-  
7 cating client-requested data for the storage operating system to retrieve from a file, di-  
8 rectory, vdisk or lun stored in the storage system;  
9 determining whether the received client read request matches any of a plurality of  
10 readset data structures ("readsets") allocated for the file, directory, vdisk or lun contain-  
11 ing the client-requested data; and  
12 performing readahead operations in accordance with a set of readahead metadata  
13 stored in a readset that is determined to match the received client read request.